CS 134: Networks | Lecture 5 **outline**, Monday, 2/6
Prof. Yaron Singer | Spring 2017

1 Overview

As a function of d, what is the fraction of nodes in the network which have degree d?

2 Normal Distributions

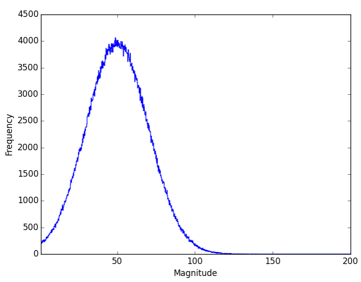

Figure 1: Density

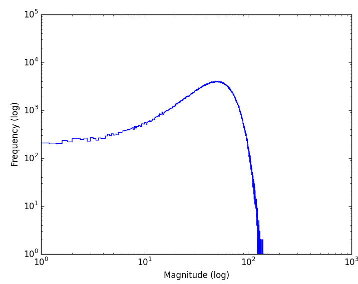

Figure 2: Density (log-log scale)

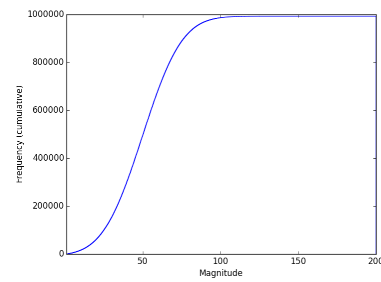

Figure 3: CDF

3 Power Law Distributions

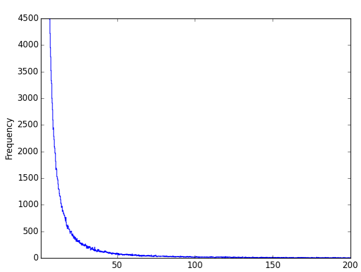

Figure 4: Density

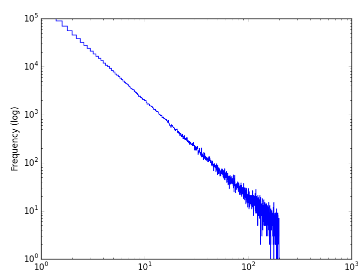

Figure 5: Density (log-log scale)

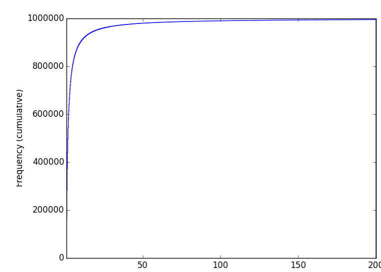

Figure 6: CDF

Definition. *A non-negative random variable X has a power law distribution if there exist constants $c, \alpha > 0$ such that:*

$$Pr[X \geq x] = c \cdot x^{-\alpha};$$

In the case where X is a discrete random variable, then there exist constants $c, \alpha > 0$ s.t.:

$$Pr[X = x] = c \cdot x^{-\alpha}.$$

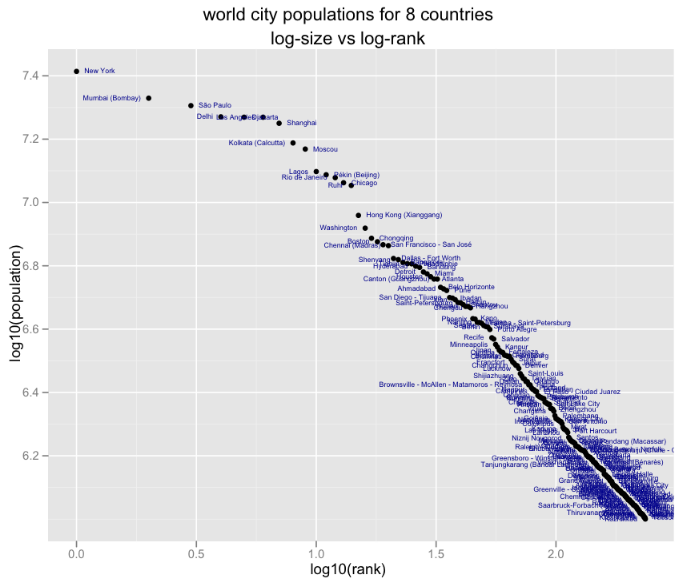

Figure 7: World city populations for 8 countries.

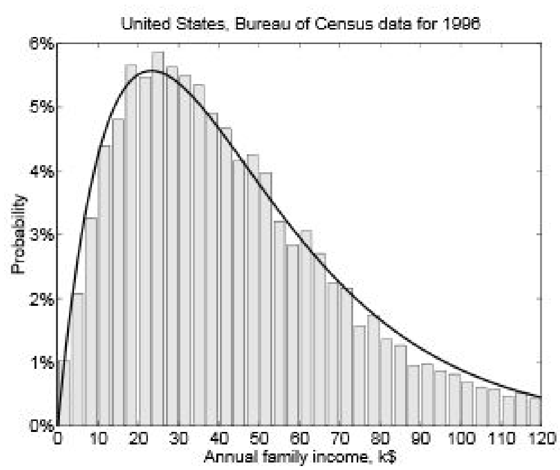

Figure 8: Income distribution in the United States.

4 Justification of Power Law in Language

Suppose that we want to design a language that optimizes the average amount of information per unit transmission.

- A language has n words and d letters.
- The jth most used word is used with probability p_j.
- The costs of sending the jth most used word is $c_j = \log_d j$.

The average cost C of transmission is then

$$C = \sum_{j=1}^{n} p_j c_j = \sum_{j=1}^{n} p_j \log_d j.$$

The average information per word is measured using the entropy H,

$$H = -\sum_{i=1}^{n} p_j \log_2 p_j.$$

To optimize the average amount of information per unit transmission cost, we wish to pick the probabilities p_j that maximize

$$A := \frac{H}{C}.$$

By taking the derivative, we get

$$\frac{\partial A}{\partial p_j} = \frac{c_j \cdot H + C \cdot \log_2(e \cdot p_j)}{H^2},$$

which is equal to zero, and therefore optimized, at

$$p_j = \frac{2^{-H \cdot c_j / C}}{e}.$$

Since $c_j = \log_d j$, we obtain that the optimal p_js follow a power law.